Rule WLM101:   Service Class did not achieve average response goal

Finding: CPExpert has detected that a service class period did not achieve the average response goal that was specified in the Service Policy in effect. | This finding applies to performance goals that specify **average response** | **time** as the performance goal.

Impact: This finding can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on performance of your computer system. The impact depends upon the importance of the service class that missed its performance goal, | and on how seriously the goal was missed.

Logic flow: This is a basic finding. There are no predecessor rules.

Discussion: The System Resources Manager (SRM) accounts for each transaction executing in the system. When the transaction ends, the SRM counts the transaction and determines the transaction's response time[1]. The SRM sums the response times for transactions ending in a service class period as each transaction ends.

The Workload Manager periodically[2] divides the sum of response times by the number of ending transactions. The result is the average response time of all transactions ending in the service class period during the previous interval.

The Workload Manager periodically assesses the performance of each service class, comparing the performance achieved by the service class against the performance goals specified for the service class[3]. This assessment is referred to as the "policy adjustment" interval, in that the Workload Manager decides whether to adjust resource policies based on whether service classes are meeting performance goals.

For service classes that have an **average response time goal**, the | Workload Manager determines whether the average response time achieved by transactions ending in the service class period is greater than the performance goal. If the average response time is greater than the

---

[1]This response time applies only to the time the transaction was in the system; it does not apply to response time delays experienced in the network.

[2]The Workload Manager computes the average transaction response time every 10 seconds, during the "policy adjustment" interval.

[3]Please see Section 4 for a more detailed description of this process.

---

performance goal, the system is not meeting performance goals for the service class period. If the importance of the service class is sufficiently high, the Workload Manager may re-allocate system resources in an attempt to meet performance goals.

At a different period (typically every 15 minutes), the SRM provides RMF with measurement data, including the elapsed and execution times of transactions ending in each service class period, and the number of transactions ending in each service class period. This information is collected by RMF and written to the SMF data set as Type 72 records. The interval at which RMF collects data and writes records typically is referred to as the *RMF measurement interval*.

RMF does not include in Type 72 records the number of instances in which any service class period did not achieve its average response goal. RMF records the total elapsed time and the number of ending transactions.

For response goals, RMF also records in Type 72 records a count of transactions that completed in varying percentages of the response goal. | These transaction counts are recorded by RMF as the "Response Time Distribution Count Table" contained in SMF Type 72(Subtype 3) records. See Rule WLM102 or Rule WLM105 for a discussion of percentile response performance goals.

The count of transactions completing in varying percentages of the performance goal is useful for analyzing performance of service classes that have a "percentile goal" specified for a service class. However, these | counts are not useful in computing average response times.

CPExpert analyzes the SMF Type 72 records to determine whether service class periods met their performance goals during each RMF measurement interval.   For service class periods that have an average response | performance goal specified, CPExpert accomplishes this simply by dividing the **number** of transactions ending in the service class (R723CRCP) into the **elapsed time** of ending transactions (R723CTET). The result is the average transaction response time **over the entire RMF measurement interval**.

CPExpert compares the average transaction response time over the entire RMF measurement interval against the performance goal specified for the service class period. If the average transaction response time is greater than the performance goal, CPExpert can conclude that the service class period did not achieve its performance goal for the RMF measurement interval. **This conclusion reveals a persistent problem**.

Some transactions executing in the service class period may have missed their performance goals, and this situation is to be expected when an average response goal is specified to the Workload Manager. The average response goal simply applies to the *average* response time achieved, which implies that the response time of some transactions may be significantly *less* than the goal and others may be significantly *more* than the goal.

It is important to appreciate that the average response time goal may not be met during a number of Workload Manager policy adjustment intervals. This circumstance may not be detected when CPExpert analyzes RMF data as described above, as the averages are computed based on an entire RMF measurement interval. CPExpert will detect a **persistent** problem, but cannot detect **periodic** problems with average transaction response times being greater than the performance goal[4].

CPExpert produces Rule WLM101 when CPExpert detects that a service class period did not meet its average response goal for an entire RMF measurement interval. CPExpert reports the total transactions that ended during the interval, the average response achieved by the transactions, and the primary and secondary causes of response delay. Additionally, CPExpert computes the contribution that the primary and secondary causes of delay made to the average transaction response time.

For example, suppose that a 100 millisecond average response time had been specified as the performance goal for a service class period serving interactive TSO transactions. CPExpert might detect that the average TSO response time was 350 milliseconds; the performance goal was missed by 250 milliseconds! CPExpert would report the number of transactions and their average response time.

CPExpert would analyze the causes of delay to TSO transactions and report the primary and secondary causes of delay. CPExpert might compute that the primary cause of delay to TSO transactions was that they were denied access to a processor for 35% of their active time, and that they were waiting for "unknown" causes[5] for another 30% of their active time.

CPExpert would report both these causes, and their respective percentages in Rule WLM101. CPExpert would continue analysis to assess which

---

[4]The Workload Manager does provide another category of service goal (the Percentile Goal) by which users can specify the percentage of transactions that should achieve their service goals. As mentioned earlier, the Percentile Goal is described in Rule WLM102 and Rule WLM105.

[5]Recall from Section 4 that the "unknown" cause is unknown as far as the Workload Manager is concerned. The Workload Manager identifies causes of delay only for those categories over which the SRM has control. Delays over which the SRM has no control are grouped together into an "unknown" category. These delays typically are certain categories of I/O delay, ENQ delay, waiting for cross-memory services, etc.

Rule WLM101.3

service classes might deprive TSO transactions from access to a processor and to assess the likely causes of "unknown" delays.

CPExpert analyzes the following possible delays to response time[6]:

- **CPU Using delay**

- **Denied CPU delay**

- **CPU Capping delay**

- **Swap-in delay**

- **MPL delay**

- **Page-in delay**

- **Non-paging DASD delay**

- **Non-DASD delay**

- **Queue delay**

- **Unknown delay**

The above causes of delay are analyzed by CPExpert in other rules.

For the purposes of identifying primary and secondary causes of response delay, CPExpert combines all auxiliary storage page-in delays into "page-in delay" to reflect the impact of auxiliary storage on response.

Additionally, CPExpert computes the average Performance Index for the service class during any measurement interval in which the performance goal was not achieved. The Performance Index is computed as the **actual response** divided by the performance **goal**.

The Performance Index gives an indication of how seriously the performance goal was missed: a Performance Index of less than 1 indicates that response was less than the performance goal; a Performance of greater than 1 indicates that response was worse than the performance goal.

The following example illustrates the output from Rule WLM101:

_____

[6]Please see Section 4 (Chapter 3.3) for a description of these delays.

Rule WLM101.4

```
    RULE WLM101:   SERVICE CLASS DID NOT ACHIEVE AVERAGE RESPONSE GOAL

       Service Class TSO (Period 1) did not achieve its response goal
       during the measurement intervals shown below.  The response goal was
       0.040 second average response, with an importance level of 2.   The
       percentages with the primary/secondary causes of delay are computed as
       a function of the average address space active time.


                               ----LOCAL SYSTEM----
                               TOTAL  AVERAGE  PERF    PLEX  PRIMARY,SECONDARY
       MEASUREMENT INTERVAL    TRANS RESPONSE  INDX     PI   CAUSES OF DELAY
       13:17-13:22,21JUN1994   5,750   0.055   1.39    1.04  DENIED CPU(51%),UNKNOWN(29%)
       13:22-13:27,21JUN1994   5,829   0.045   1.12    1.02  UNKNOWN(40%),DENIED CPU(36%)
```

The information associated with Rule WLM101 is shown based on data
collected by the *local system,* which is the system being analyzed for
performance purposes.

CPExpert also computes and reports a *sysplex* Performance Index. The
WLM maintains both a "sysplex Performance Index" and a "local system
Performance Index."   Briefly, the WLM first examines the sysplex
Performance Index to determine whether a service class period is missing
its performance goal and whether action should be taken. After the sysplex
Performance Index is examined at a particular Goal Importance level, the
WLM then examines the local system Performance Index. Rule WLM140
explains this WLM logic in more detail, and describes the implications of the
WLM logic.

Suggestion: There are no suggestions with this finding. CPExpert will continue analysis
and other rules will be produced to provide more information.

Rule WLM101.5